Music Piracy: Bad for Record Sales but Good for the iPod?

Leung, Tin Cheuk

Chinese University of Hong Kong

2012

Music Piracy:
Bad for Record Sales but Good for the iPod?

Tin Cheuk Leung[*]

Chinese University of Hong Kong

October 2, 2012

Abstract

Music piracy is a double-edged sword for the music industry. On the one hand, it hurts record sales. On the other hand, it increases sales of its complements. To quantify the effect of music piracy, I construct a unique survey data set and use a Bayesian method to estimate the demand for music and iPods, and find three things. First, music piracy decreases music sales by 24% to 42%. Second, music piracy contributes 12% to iPod sales. Finally, counterfactual experiments show that Apple's revenue could increase by $36 per student if music were free.

1 Introduction

The belief that music piracy hurts record sales prevails in the music industry.[1] In 1999, record sales started to decline after more than a decade of steady growth. The very same year, Napster, the first peer-to-peer (P2P) software used to pirate music, began operations, which music industry representatives, such as the Recording Industry Association of America (RIAA) and the International Federation of the Phonographic Industry (IFPI), argue is no coincidence. A number of economists have combined data on illegal downloading from a variety of sources with data on album sales and found that music piracy has led to a decline in the latter (see Blackburn (2004), Liebowitz (2006), Peitz and Waelbroeck (2004), Rob and Waldfogel (2006) and Zentner (2006).) Oberholzer-Gee and Strumpf (2007) disagree, however, arguing that music piracy "allows users to learn about music they would not otherwise be exposed to" and thus may actually boost record sales. They construct a unique data set using weekly volumes of illegal downloads and show that music piracy's effect on record sales " is not statistically distinguishable from zero. The economic effect of the point estimate is also small." These contradictory findings cast doubt on the music industry's belief in the damaging effects of music piracy. Whether music piracy hurts record sales remains an open empirical question.

Although, as noted, many economists argue that music piracy hurts record sales, Oberholzer-Gee and Strumpf (2010) argue that it can increase revenue from music complements such as concerts. When music piracy drives the effective price of an album down to nearly zero, more consumers become familiar with the artist's music, thus driving up demand for concerts. Mortimer, Nosko, and Sorensen (2012) show empirically that music piracy has a significant effect on concert demand, particularly for small artists. Concerts are not the only music complements to benefit from music piracy. Figure 1 suggests that the iPod, the most popular MP3 player by far, has gained considerably from such piracy. Although record sales have declined by more than 25% since Napster was launched, Apple has seen exponential growth in iPod sales since the device was introduced in 2001. Revenue from iPod sales grew from $344 million in 2003 to $7.6 billion in 2006. If there is strong complementarity between music piracy and iPod sales, then Netanel (2003) and Fisher (2004) suggest that it might be welfare-improving to have a regime

---

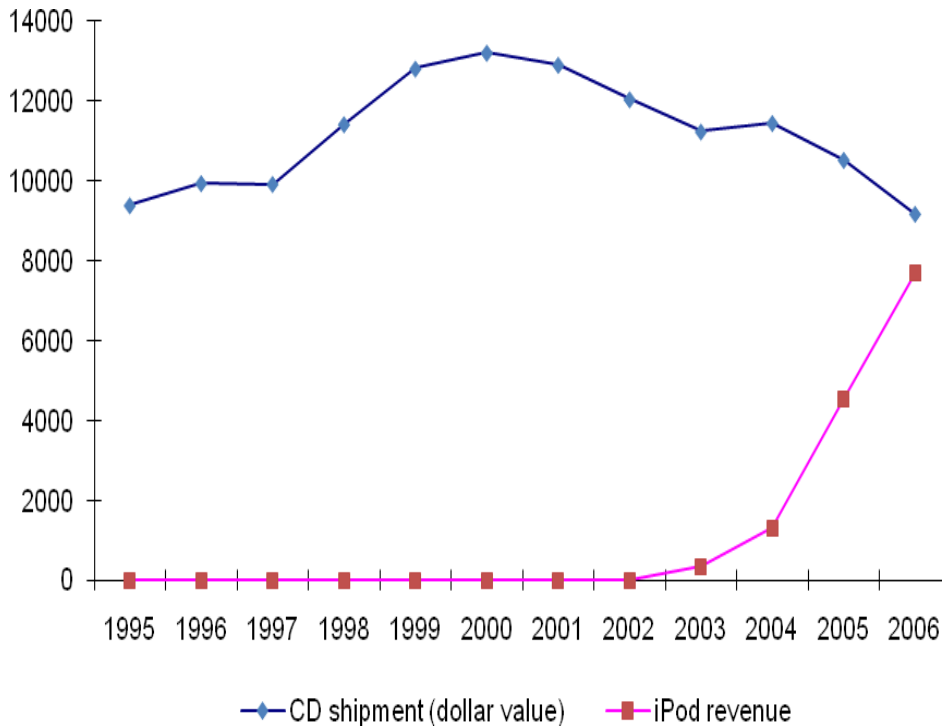[1]I use music piracy and illegal downloading interchangeably in this paper.

Figure 1:
RIAA and Apple Inc.: CDs Revenue Decreasing, iPods Revenue Growing in the U.S. ($ millions)

in which music piracy is legal and a government-financed fund compensates music producers according to the download rates of their records. How much music piracy benefits the iPod sales is another open and important empirical question.

To answer these two empirical questions, I constructed a unique conjoint data set derived from a survey of 884 University of Minnesota undergraduates. The students were first asked to report their demographic information and their recent consumption of both music and iPods. Then, in the conjoint survey, they were asked to make choices about music (from both legal and illegal sources) and iPods in 12 hypothetical situations. Green and Rao (1971) introduced conjoint survey analysis as a way to elicit demand estimates. Conjoint survey data are also known as stated-preference data, as opposed to revealed-preference data collected from real world observations. There are two main advantages to the use of conjoint survey data rather than real market data in this research. First, a conjoint survey is possibly the only way to create a panel data set on the consumption of illegal downloads, legal music, and iPods. Because it

3

is essential to ascertain the influence of any copyright regime changes in the music industry on related products such as iPods, I require a clean panel data set on the consumption of both music (from legal and illegal sources) and iPods. Second, use of a conjoint survey allowed the use of instruments for illegal downloads that are unavailable in other studies for the reasons discussed in Section 3.

Several studies argue that conjoint survey data generate reliable demand estimates,[2] and applications of conjoint survey analysis abound. For example, Leung (2012) estimates the substitutability of street and Internet piracy of Microsoft Office in Hong Kong, and Hensher and Louviere (1983) forecast the choice of attendance at various types of international expositions. Hensher (1994) reviews the development of conjoint analysis's use in estimating transportation choice. Many multinational corporations, including Marriott, Procter & Gamble (P&G), and General Motors, also use conjoint survey data to estimate the demand for new products (see Green, Krieger, and Wind (2004) and Orme (2005)).

My empirical analysis consists of three parts. First, I set up a demand system with three types of music: CDs, legally purchased songs from iTunes, and pirated songs from P2P websites. I estimate this system of three simultaneous equations using the three-stage least-squares method. My results contrast with those of Oberholzer-Gee and Strumpf (2007), and support the belief that music piracy hurts record sales. Second, I use the estimates from the first part to establish a random-coefficient discrete-demand model for iPods. I follow Rossi, Allenby, and McCulloch (2005) in setting up a hierarchical Bayesian discrete-demand model, with a mixture of normal priors, and then use a hybrid of Gibbs Sampling and the Metropolis-Hasting algorithm to implement posterior inference. These estimates indicate that music piracy boosts demand for and sales of iPods. Third, I use the estimates from the first and second parts of the analysis to conduct counterfactuals to predict the changes in demand for music and iPods in different regimes. The results show that approximately 12% of Apple's revenue comes from music piracy, which translates into $1.1 billion using 2008 revenue figures.

The remainder of the paper is organized as follows. Section 2 briefly describes the current situation of growing music piracy and declining record sales. Section 3 discusses the conjoint

---

[2]Carlsson and Martinsson (2001) and Hensher, Louviere, and Swait (1999) collect both stated-preference and revealed-preference data on donation choice and freight shipper choice, respectively, and show that the hypothesis of parameter equality holds for most parameters across the two data sources.

survey data set. Sections 4 and 5 describe the set-up of the demand estimates for music iPods, respectively, and discuss the estimation results. Section 6 reports the results of counterfactual experiments using the results from Sections 4 and 5. Section 7 concludes the paper.

2 Music Piracy Growing, Record Sales Shrinking

2.1 *Music Piracy is Growing*

P2P technology has enhanced the ease and speed with which music can be pirated, and thus triggered the growth of music piracy. In 1999, the first P2P software, Napster, was introduced, and the number of music pirates has been growing ever since.
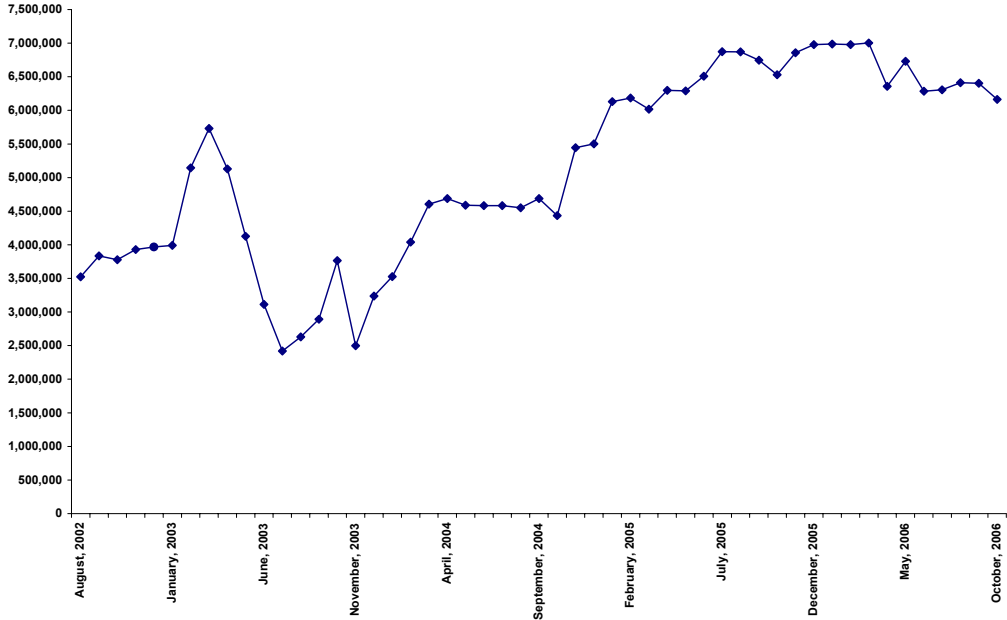
Figure 2:
Big Champagne: Avg Simultaneous P2P Users in the U.S. Has Been Growing

People download music illegally from the Internet because the cost of doing so is low. If the cost were even lower, even more people would engage in music piracy. Recent advances in Internet connection speeds has reduced the time cost of pirating music over the Internet, which has led to the further growth of music piracy. The marketing research firm Big Champagne

reports an increasing trend in searching, clicking, and pirating music. Figure 2 shows that the average number of simultaneous users of P2P software in the U.S. increased from 3.5 million in August 2002 to more than 6 million in October 2006.[3] This growing number of music pirates translates into a huge number of pirated songs. The IFPI estimates that almost 20 billion songs were illegally downloaded in 2005.

2.2   Record Sales are Shrinking

Music is important to Americans. The average American enjoys almost an hour of music every day.[4] Before Napster, a major source of such enjoyment was records. Record sales almost quadrupled between 1990 and 1999. As we can see from Figure 1, the 1990s were the heyday of the music industry. However, since Napster appeared on the scene in 1999, record sales have declined by \$3.6 billion.

The music industry believes that music piracy hurts record sales and has taken action based on that belief. In 1999, the RIAA sued Napster, leading to the latter's demise in 2001. In addition, between 2003 and 2005, the RIAA sued approximately 11,700 individual music pirates, despite the reputation cost of effectively suing its own customers (Associated Press (2005)).

Economists have created a variety of illegal download data sets to estimate the effect of music piracy on record sales. Rob and Waldfogel (2006), for example, conducted a survey in a number of universities to compile a panel data set on both illegal downloads and album consumption. Oberholzer-Gee and Strumpf (2007) and Blackburn (2004) created music piracy panel data sets by tracking individual illegal downloading behavior on P2P software. All of these researchers supplement their data with aggregate record sales data from the RIAA or Nielson Soundscan.

Both Oberholzer-Gee and Strumpf (2007) and Rob and Waldfogel (2006) run a regression of the following form to determine the displacement effect of illegal downloads on album sales.

$$A_{jt} = X_{jt}\beta + \alpha D_{jt} + \epsilon_{jt}, \qquad (1)$$

---

[3]There was a wave of lawsuits against individual pirates in 2003, which caused a decline in the number of P2P users at the time. This fact motivates me to include the expected punishment as one of the covariates in the conjoint survey as discussed in Section 3.

[4]See Table No.909, "Media Usage and Consumer Spending: 1993 to 2003," in the 2000 U.S. Statistical Abstract.

where A_{jt} is the sales of album j at time t, D_{jt} is the number of illegal downloads, and X_{jt} are other covariates. D_{jt} may be endogenous. Popular albums usually attract more downloads, in which D_{jt} will be positively correlated with ϵ_{jt}, and the estimate of α will have an upward bias. The aforementioned researchers deal with this problem by finding instruments for illegal downloads that are not themselves related to album sales and thus not correlated with ϵ_{jt}. Oberholzer-Gee and Strumpf (2007) use the number of German high school students on vacation as such an instrument based on the premise that German high school students spend more time pirating music during their holidays.[5] Rob and Waldfogel (2006) use the speed of students' Internet connections as the instrument based on the assumption that students do not choose an Internet speed based on their music preferences.

Rob and Waldfogel (2006) and Blackburn (2004) find that music piracy hurts record sales. Rob and Waldfogel (2006) find that "one (illegally) downloaded album reduces music purchases [by] roughly one-fifth of an album." Blackburn (2004) estimates in his counterfactuals that "the lawsuits brought by the RIAA have resulted in an increase in album sales of approximately 2.9% during the 23 week period after the lawsuit strategy was publicly announced."

Table 1: US Legal Digital Music Market Growing (millions)

	2004	2005	2006
Broadband lines	34	43	57
Single track downloadeds	143	353	582
Album downloads	6	16	33
Mobile subscriptions	163	174	194

Sources: IFPI "Digital Music Report" 2006 and 2007.

Oberholzer-Gee and Strumpf (2007), however, argue that music piracy's effect on record sales "is not statistically distinguishable from zero. The economic effect of the point estimate is also small." They argue that there may be other more important factors leading to the decline in record sales. First, we might be seeing a shift in entertainment spending from recorded music toward recorded movies. Second, many people may have replaced their old LPs with CDs in the mid-1990s, thus boosting record sales in that period. By 1999, coincidentally the year that Napster began to operate, however, they had completed their replacement process. Third, the

---

[5]Liebowitz (2007) points out that Oberholzer-Gee and Strumpf (2007) make a contradictory claim in their quasi-experiment stating that illegal downloading decreases in the summer because American college students lose their broadband connections during their vacation. As both countries have both high school and college students, theoretically there should be no difference in the way that school holidays affect illegal downloading.

emergence of digital (online) music stores such as iTunes provides an even closer substitute for CDs. Table 1 shows that the number of legal downloads of both single tracks and albums increased by more than 50% per year from 2004 to 2006. In addition to these other factors, Oberholzer-Gee and Strumpf (2007) argue that music piracy may in fact boost record sales because it allows consumers to learn about music they would not otherwise be exposed to.

These conflicting findings lead to my first question: Does music piracy hurt record sales? If so, by how much? To answer the question, I use a different approach from that of previous researches. Section 3 describes the conjoint survey data set and compares the pros and cons of this data set with these of the data sets used by others.

3 Data Collection and Description

Table 2: Percentage of U.S. Adult Population Answering YES to "Do you ever download music files on your computer so you can play them at any time you want?"

	Aug-Sep 2001	Oct 2002	June 2003	Nov 2003	May-June 2004	Feb 2005
Overall	15	19	19	9	13	13
18-29	36	41	43	23	31	32
30-49	16	21	20	9	11	13
50-64	6	8	8	4	6	7
65+	2	3	1	2	2	1
Men	19	22	23	12	17	14
Women	13	16	15	7	9	12

Source: Pew Internet Project.

I collected conjoint survey data from college students, who generally have a lower income and greater exposure to the Internet than other age groups. As Table 2 shows, college students also tend to download or pirate more music than other age groups.

3.1 *Conjoint Survey*

I conducted the survey in Fall 2007 and Spring 2008 among seven undergraduate classes. Of the approximately 1800 students registered for these classes, 884 attended class on the day of the survey and turned in completed surveys.

The survey focuses on one dominant brand of MP3 player, the Apple iPod, because Apple

Table 3: Apple Dominates the MP3 Market

Brand	Unit Share
Apple	72.7%
Sandisk	8.9%
Microsoft	3.2%
Creative Labs	2.9%
Samsung	2.0%

Source: NPD Group.

dominates the MP3 market. As Table 3 shows, Apple enjoys more than 70% share of the MP3 market. Sandisk, its closest competitor of Apple, sells only one-eighth of what Apple does.[6]

The survey comprised of two parts. The first asked students to report demographic information and give details concerning their Internet access and music and iPod consumption preferences. The second was the conjoint survey. Green and Rao (1971) were the first to introduce conjoint analysis to the marketing field. I followed the approach of Louviere and Woodworth (1983) in using choice-based conjoint analysis, which integrates conjoint analysis and discrete choice analysis. The questions in conjoint surveys are not descriptive such as "How much would you be willing to pay for an iPod?" Instead, they ask students to make concrete choices such as "Given brand A, B, and C with different attributes and prices, which one would you buy?" Conjoint survey data are also known as stated-preference data, as opposed to revealed-preference data, which are collected from real market transactions.

The conjoint survey used in this study included 12 hypothetical tasks. In each task, the respondents were given the option of listening to music on an iPod, a computer or a radio (which I treat as an outside choice). The choices differed in the level of each of the following six choice-specific covariates.

- Price of an iPod (varies from $30 to $650)

- Capacity of an iPod (varies from 1 gigabyte to 8 gigabytes)

- Probability of getting caught pirating music (varies from 0 to 1)

- Fine per song if caught pirating music (varies from $0 to $10,000)

---

[6]The respondents also showed a distinct preference for iPods in a trial run of the survey that included other brands of MP3 players.

- Price per song in iTunes (varies from $0.1 to $3)

- Price per CD (varies from $1 to $30)

There were five to ten levels for each covariate within the pre-specified range.

I followed the following three principles proposed by Sawtooth Software (2008) to draw the levels of each covariate.

1. **Minimal Overlap**: Each covariate level is shown as few times as possible in a single task.

2. **Level Balance**: Each covariate level is shown an approximately equal number of times.

3. **Orthogonality**: Covariate levels are chosen independently of other attribute levels, such that each level's effect on utility can be measured independently of all other effects.

The student were asked to complete two sub-tasks in each of the 12 tasks. In the first sub-task, they were asked to imagine that they did not have an iPod and then to rank the three choices in the task. Figure 3 presents a sample of the first sub-task.

Before ranking the choices, students knew roughly what their level of music consumption would be under each. For instance, the students who ranked the iPod as their top choice were also those who estimated that they would buy or pirate a considerable amount of music. In the second sub-task, I assigned the students one of two choices—iPod or computer. This assignment may or may not have been their first choice in the first sub-task. Given their assigned choice and associated music prices, students were then asked how they would change their music consumption from the previous month and estimate their consumption from three sources (CDs, iTunes and P2P websites). Figure 4 shows a sample of the second sub-task, which was included in the conjoint survey administered to two of the seven classes, meaning that 270 students completed it.

3.2   *Conjoint Survey Data versus Real Market Data*

There are several advantages and disadvantages to using conjoint survey rather than real market transaction data to estimate the demand for music and iPods.

When you listen to music,
Your first choice is: _____   Second choice is: _____

Option 1	Option 2	Option 3
iPod nano	**Your computer**	**Radio**
$US200, 4GB		
Free P2P downloading **illegal**. Fine: $US200/song downloaded that month Chance of getting caught per month: **1 in 2000 songs**	Free P2P downloading **illegal**. Fine: $US50/song downloaded that month Chance of getting caught per month: **1 in 100,000 songs**	
iTunes: **$ US 0.3/song** CD: **$5 each**	iTunes: **$ US 3/song** CD: **$10 each**	Free music

Figure 3:
A Sample of the First Sub-Task

Suppose you had an iPod last month and prices of music from different sources were as follows:

P2P downloading	Illegal	Fine: $200/song	Prob. Of getting caught: 1 in 2000 songs
iTunes	$0.3/song		
CDs	$5/CD		

What would be your music consumption from the 3 sources?

P2P downloading   : _____ songs
iTunes   : _____songs
CDs   : _____CDs

Figure 4:
A Sample of the Second Sub-Task

One of the disadvantages of doing so is that conjoint analysis requires attributes to be quantifiable, and some attributes, such as the quality of sound from an iPod and a computer, are difficult to quantify. However, because these attributes are unlikely to vary across the different copyright regimes I discuss in the counterfactuals section, omitting them is unlikely to affect the results.

Some researchers have also expressed concerns about the validity of conjoint survey data. Some consider real market data to be more reliable because it is revealed-preference data. However, since Green and Rao (1971) introduced conjoint survey analysis to the marketing area, it has been widely adopted in the marketing literature to elicit demand estimates. As mentioned in the introduction, applications of such analysis abound. Several studies also argue that conjoint survey data can generate reliable demand estimates.

Among the advantages of using conjoint survey data rather than real market data in this research were the following. First, conjoint survey was possibly the only way to create a panel data set encompassing the consumption of illegal downloads, legal music, and iPods. Oberholzer-Gee and Strumpf (2007) and Blackburn (2004) gather panel data sets on music piracy by tracking individual illegal downloading behavior on a P2P network. They then combine weekly album sales with their novel data on weekly download volumes to estimate the effect of illegal downloads on album sales. Rob and Waldfogel (2006) carry out surveys among college students to create a panel data set incorporating legal music consumption and illegal downloading behavior, and then use their data set to estimate effect of piracy on record sales. However, as I argued before, it is important to ascertain the effect of any copyright regime changes in the music industry on other related products such as iPods. Doing so requires a clean panel data set on both the consumption of music (both legal and illegal) and iPods. To the best of my knowledge, this paper is the first to construct such a panel data set using a conjoint survey.

Second, conjoint survey analysis provides good instruments. Both Oberholzer-Gee and Strumpf (2007) and Rob and Waldfogel (2006) use an instrumental variable approach to deal with endogeneity in Equation (1). In this paper, I use the expected probability of getting caught and possible fines as instruments for illegal downloads. Although each affects illegal downloads, neither has a direct effect on legal music consumption.[7] However, despite serving as instruments

---

[7]I also used the price of an iTunes song, the price of a CD, and the price of an iPod as instruments for their

for my purposes here, they would not work in other cases for two reasons. First, there is greater variation in expected punishment in my conjoint survey than in the real world. In a conjoint survey, the designer can vary the prices of different choices within a pre-specified range. For instance, I vary the fine per song from $0 to $10,000, whereas the real fine is usually $10 to $50 per song. Basic econometrics tells us that a larger degree of variation in independent variables (the expected probability of getting caught and possible fines in our case) provides more information about their effect on the dependent variable (illegal downloads here). The second reason is that the levels of these two covariates are drawn exogenously and independently owing to the orthogonality principle described in the previous subsection. Hence, they do not correlate with ϵ_{jt} in Equation (1) and can serve as instruments for illegal downloads.

3.3 Data Description

Completed surveys were received from 884 students. Most were typical university students: They did not have a high income, with around 90% of them earning less than $200 per week. In addition, they reported spending an average of three to four hours a day on the Internet.

Table 4: Summary Statistics of the Data

	Mean (s.d.)	Min	Max
Age	18.94 (1.87)	13	45
Owns an MP3 player	0.86 (0.35)	0	1
Owns an iPod	0.72 (0.45)	0	1
Number of songs on computer	2508 (4773)	0	75000
Bought a CD in past month	0.28 (0.45)	0	1
Bought songs from iTunes last month	0.32 (0.47)	0	1
Downloaded free songs last month	0.54 (0.50)	0	1

N=844.

Table 4 shows that the respondents expressed great interest in listening to music. They reported having an average of 2508 songs on their computers. They both bought music, and pirated it. Twenty-eight percent and 32% had bought a CD or a song from iTunes, respectively, in the previous month. In addition, more than half (54%) of them had downloaded songs illegally from the Internet in the previous month.

corresponding demands.

13

More than 70% of the students said they owned an iPod.[8] Unsurprisingly, the students who owned more music were more likely to own an iPod. If I define students who have more than 1000 songs on their computers as music lovers and others as non-music lovers, 80% of music lovers own an iPod, whereas only 60% of non-music lovers do.

Table 5: Summary Statistics of Conjoint Survey

Choice	First Choice (%)	Second Choice (%)
iPod	46.11%	28.50%
Computer	36.63%	37.23%
Radio	17.27%	34.28%

N=10608.

Table 5 shows that in the first sub-task of the conjoint survey, iPod was the most popular of the three choices, with computer coming second.

An iPod was students' first choice almost half of the time. Not only did they show a preference for iPods, but they also prefered lower prices. Table 6 shows that when iPod or music prices decreased, iPods were the first choice more often.

Table 6: iPod as First Choice under Extreme Attribute Levels

Attribute	Lowest level	Highest level
iPod price	56.75%	42.21%
Price per song in iTunes	74.22%	42.37%
Price per CD	66.88%	39.66%
Probability of getting caught	85%	23.89%
Fine per song	85%	39.22%
GB	53.04%	53.51%

N=10608.

As noted, I included the second sub-task in the conjoint survey in two of the seven classes. Of the approximately 884 students surveyed, 270 completed surveys with the second sub-task. These 270 students have similar characteristics to the overall sample.

4 Music Demand

I used the data from the second sub-task in the conjoint survey to estimate music demand. The three dependent variables are CDs, iTunes songs, and pirated songs from P2P websites, and the

---

[8]Among the students who owned an MP3 player, more than 80% owned an iPod, which justifies my focus on iPods specifically, instead of MP3 players generally, in this paper.

independent variables are the prices of music from different sources and demographic variables.

As my first question in this study was whether music piracy hurts record sales, I needed to determine the effect of pirated songs on CD and iTunes song consumption. The problem can be expressed as a simultaneous equations problem. The simultaneous demands for music from agent i in task t are

$$\log Y^*_{itp} = \mathbf{z}'_{itp}\gamma_p + \log(Y^*_{its})\phi_{ps} + \log(Y^*_{itc})\phi_{pc} + u_{itg} \tag{2}$$

$$\log Y^*_{its} = \mathbf{z}'_{its}\gamma_s + \log(Y^*_{itp})\phi_{sp} + \log(Y^*_{itc})\phi_{sc} + u_{its} \tag{3}$$

$$\log Y^*_{itc} = \mathbf{z}'_{itc}\gamma_c + \log(Y^*_{its})\phi_{cs} + \log(Y^*_{itp})\phi_{cp} + u_{itc} \tag{4}$$

where subscripts p, s and c denote P2P (pirated songs), iTunes songs, and CDs. For $g \in \{p, s, c\}$ $Y^*_g = Y_g + 1$, Y_g is the consumption of g; $\mathbf{z}_g$ is a vector of exogenous regressors, including prices, that are uncorrelated with u_g; and u_{itg} are i.i.d. over i and t and homoskedastic, but correlated across g. Table 7 presents all of the $\mathbf{z}$.

Table 7: Exogenous Regressors in Music Demand

$\pi^* = \pi + 0.001$	probability of getting caught
$f^* = f + 1$	fine per song if caught
P^*_s	price per iTunes song
P^*_c	price per CD
$iPod$	indicator for iPod
$h^* = h + 1$	hours spent on Internet per day
$Antivirus$	indicator for having antivirus software
$P2Pfd$	indicator for having friend who engages in piracy
$Dorm$	indicator for living in dorm
$Income$	level of income
$Prob$	perceived probability of getting caught in real world
$MusInt$	level of music interest
$P2P^*$	illegal songs downloaded last month
$iTunes^*$	iTunes songs purchased last month
CD^*	CD purchased last month

Each dependent variable has its own instruments. For instance, the probability of getting caught pirating music (π) instruments for the demand for pirated music from P2P websites, the price per song in iTunes instruments for the demand for iTunes songs; and the price per CD instruments for CD demand. I used the three-stage least-square method to estimate this simultaneous equations system, and the results are presented in Table 8.

There are several things to note. First, the students pirated more music when they owned

Table 8: Music Demand/Month (std. err.)

	$\log(P2P+1)$ Demand/Month	$\log(iTunes+1)$ Demand/Month	$\log(CD+1)$ Demand/Month
Constant	0.27 (0.19)	2.21 (0.07)	0.93 (0.04)
$\log Y_p^*$		-0.07 (0.01)	-0.04(0.01)
$\log Y_s^*$	-0.33 (0.03)		-0.09 (0.01)
$\log Y_c^*$	-0.02 (0.10)	-0.09 (0.06)	
$iPod$	0.37 (0.06)	0.14 (0.03)	
$\log \pi^*$	-0.32 (0.01)		
$\log f^*$	-0.21 (0.01)		
$\log P_s^*$		-1.82 (0.04)	
$\log P_c^*$			-0.28 (0.01)
$Antivirus$	0.29 (0.12)		
$\log h^*$	-0.20 (0.07)		
$P2Pfd$	0.06 (0.03)		
$Dorm$	0.11 (0.06)		
$Income$	0.05 (0.03)		
$Prob$	0.05 (0.01)		
$MusInt$		0.02 (0.02)	0.02 (0.01)
$MusInstr$			-0.004 (0.015)
$\log P2P^*$	0.35 (0.01)		
$\log iTunes^*$		0.23 (0.01)	
$\log CD^*$			0.25 (0.01)

N=3240.

an iPod. Second, the law of demand holds, that is, the demand for music dropped when prices increased. However, because I used a simultaneous equations system, the coefficients of price do not fully reflect the effect of price changes on all three demands equations.

Table 9 shows how the demand for music from different sources changed when music prices changed and when an iPod was not available . When students could not own an iPod, in contrast to the real world in which 72% owned an iPod, they pirated 20.21% less music from P2P websites, and consumed 7.77% fewer songs from iTunes, but consumed 1.8% more CDs.

Table 9: Percentage Change in Demand for Music with Price Changes

	y_{P2P}	y_{iTunes}	y_{CD}
$\pi(0.0001 \to 0.0002)$	-2.83%	0.20%	0.10%
f ($100 \to 200$)	-13.76%	1.03%	0.54%
$P_s(0.99 \to 1.09)$	3.05%	-8.73%	0.72%
$P_c(15 \to 16.5)$	-0.01%	0.22%	-2.51%
To a "no-iPod" world	-20.21%	-7.77%	1.80%

The probability of getting caught and the payment of a fine are significant components of the price of (or punishment for) pirating music. Students pirated less music when the punishment was more severe. When the probability of getting caught increased 100% from 0.01% to 0.02%,

students pirated 2.83% less music from P2P websites, and consumed 0.20% more songs from iTunes and 0.10% more CDs. When the fine per song increased 100% from $100 to $200, students pirated 13.76% less music from P2P websites, and consumed 1.03% more songs from iTunes and 0.54% more CDs.

Students bought fewer iTunes songs when these were more expensive. When the price per song rose 10% from $0.99 to $1.09, they bought 8.73% fewer songs from iTunes. They also pirated 3.05% more music from P2P websites and consumed 0.72% more CDs.

Similarly, student bought fewer CDs when CDs were more expensive. When the price of a CD increased 10% from $15 to $16.5, they bought 2.51% fewer CDs. At the same time, they also pirated approximately the same amount of music, but bought 0.22% more songs from iTunes.

My estimates are consistent with those of Shiller and Waldfogel (2011), who estimate the demand for iTunes songs using survey data collected from 500 students. They find that when the price of an iTunes song increases from $0.99 to $1.87, demand drops by 42% from 7434 to 4351. I identified a similar price effect on demand for iTunes songs using the estimates in Table 8. When the price of an iTunes song was increased from $0.99 to $1.87 in the survey, demand drops 49%, which is reasonably close to the 42% in Shiller and Waldfogel (2011). At the same time, I also found a price effect on the demand for other types of music. Students pirated 25% more music and bought 6% more CDs in this case.

Note that record sales from different sources are substitutes for one another. When students bought 10% more CDs, demand for iTunes songs decreased by 0.9%. At the same time, demand for CDs decreased by 0.9% when consumption of iTunes songs increased by 10%. The emergence of online music stores such as iTunes plays a role in the decline of revenue from CD sales.

Table 10: Piracy Elasticity of Sales (%)

Oberholzer-Gee and Strumpf	-0.00001
Rob and Waldfogel	-0.13
Blackburn	-0.18
This paper (CD sales)	-0.04
This paper (iTunes sales)	-0.07

Finally, it is clear that music piracy does hurt record sales.[9] When students pirated 10% more

---

[9]Table 10 reports the piracy elasticity of sales. Oberholzer-Gee and Strumpf (2007) and Rob and Waldfogel (2006) report only estimates of the displacement effect of illegal downloads (P2P) on album sales. I combine those estimates with their sample statistics on album consumption and illegal downloads to calculate the elasticities.

music through P2P websites, they bought 0.7% fewer iTunes songs and 0.4% fewer CDs. This result is both economically and statistically significant and contrary to the claim of Oberholzer-Gee and Strumpf (2007), thereby corroborating the findings of other economists. Rob and Waldfogel (2006) report that individuals buy 1.3% fewer records (including iTunes songs and CDs) when they pirate 10% more music, whereas Blackburn (2004) suggests a higher number: people buy 1.8% fewer records when they pirate 10% more music.

5 Discrete-Choice Demand for iPods

The results in Section 4 suggest that music piracy does hurt record sales. Before quantifying the welfare implications of the three copyright regimes, I here build a discrete-choice demand model for iPods to quantify the complementary relationships between music and iPods.

In each of the 12 first sub-tasks in the conjoint survey, students were asked to rank three options for listening to music: an iPod, a computer, and the radio (which I treat as an outside good). These rankings served as the students' choices, and thus constitute the dependent variables in the demand estimation.

Students would generally be expected to have a rough estimate of their music consumption before they purchased an iPod. The average lifetime of an iPod is two years, and students would only buy one if they thought they would buy or pirate a considerable amount of music in those two years. I thus entered the estimated demands for music from the previous section into the indirect utility of a choice to account for how music complements that choice.[10] The indirect utility of choice j for student i in task t is

$$U_{ijt} = \beta_{ij} + \alpha_{i,price}P_{jt} + \alpha_{i,gb}GB_{jt} + \alpha_{i,p2p}\hat{P2P}_{ijt} + \alpha_{i,itunes}\hat{iTunes}_{ijt} + \alpha_{i,cd}\hat{CD}_{ijt} + \epsilon_{ijt}, \quad (5)$$

where P_{jt} is the price of choice j in task t, GB_{jt} is the capacity (in gigabytes) of choice j in task t, and ϵ_{ijt} is the usual i.i.d. logit error.

---

[10]I did not correct the standard errors in the second-stage estimation of discrete demand. In other words, I treat the estimated demands for music as true demands.

We can express the demand parameters of student i in Equation (5) as

$$\Theta_i \equiv [\beta_{i1}; \beta_{i2}; \alpha_{i,price}; \alpha_{i,gb}; \alpha_{i,p2p}; \alpha_{i,itunes}; \alpha_{i,cd}],$$

where Θ_i is a 1×7 vector of individual parameters. Θ is a $n_i \times 7$ matrix whose ith row is Θ_i, and n_i is the number of students in the sample. Define X_{jt} as a 1×7 vector of covariates in Equation (5), and we can then express the indirect utility as

$$U_{ijt} = \Theta_i X_{jt} + \epsilon_{ijt}. \tag{6}$$

The likelihood that student i will choose j in task t takes the following logit form.

$$Pr_{ijt} = \frac{\exp(U_{ijt})}{\sum_k \exp(U_{ikt}) + 1}. \tag{7}$$

As Berry, Levinsohn, and Pakes (1995), Nevo (2000), Petrin (2002) and Rossi, Allenby, and McCulloch (2005) argue, random coefficient models generate better estimates of consumer demands than do homogenous coefficient models. In this data set, it is natural to consider that the students have heterogeneous coefficients. For instance, an average student may be more responsive to iPod price changes than an iPod lover, which translates into a higher price coefficient (in absolute value) for the average student.

I follow Rossi, Allenby, and McCulloch (2005) in using a hierarchical Bayesian model with a mixture of five components of normal priors to estimate the random coefficients. This approach is more flexible than the classical approach because it does not restrict the coefficients to coming from a normal distribution. Moreover, this approach allows for correlated coefficients without additional computation time.

Because the students provided demographic information in the survey, I include aspects of that information in the demand model to account for the observed heterogeneity across students. Define Z_i as a $1 \times n_z$ vector of the demographic characteristics of student i, where n_z is the number of such characteristics. Define Z as an $n_i \times n_z$ matrix. The demand model, in which the

unobserved heterogeneity is distributed as a K mixture of normal, can be expressed as follows.

$$
\begin{aligned}
U_{ijt} &= \Theta_i X_{jt} + \epsilon_{ijt} \\
\Theta_i &= Z_i \triangle + u_i \\
u_i &\sim N(\mu_{ind_i}, \Sigma_{ind_i}) \\
ind_i &\sim \text{Multinomial}_K(\gamma)
\end{aligned}
$$

γ is a vector giving the mixture probabilities for each of the K components. The complete specification with priors over the mixture probabilities (α), mean ($\bar{\mu}$ and a_μ^{-1}), and covariance matrices (v and V) is

$$
\begin{aligned}
\gamma &\sim Dirichlet(\alpha) \\
\mu_k | \Sigma_k &\sim N(\bar{\mu}, \Sigma_k \times a_\mu^{-1}) \\
\Sigma_k &\sim IW(v, V) \\
\{\mu_k, \Sigma_k\} &\text{ independent}
\end{aligned}
$$

5.1 *Estimation*

I follow Rossi, Allenby, and McCulloch (2005) in using a hybrid of Gibbs sampling and the Metropolis-Hasting method to implement posterior inference for this model. I use a hybrid Metropolis method that employs customized Metropolis candidate density to draw Θ_i for each student. Conditional on Θ_i, I use an unconstrained Gibbs sampler to draw δ, μ_k, and Σ_k.[11] In particular, I alternately obtain draws between the individual-level parameters in (8) and hyperparameters in (9):[12]

$$
\Theta_i | ind_i, Z_i \triangle, \mu_{ind_i}, \Sigma_{ind_i} \tag{8}
$$

$$
\gamma, ind, \triangle, \{\mu_k\}, \{\Sigma_k\} | \{\Theta\} \tag{9}
$$

---

[11] Constraints must be imposed on the Gibbs sampler to fix an identification problem called "label switching" if inference is desired for the mixture component parameters. This is not a problem here because I am interested in estimating individual student parameters and their distribution across students alone. An unconstrained Gibbs sampler is enough to ensure identification. See Rossi, Allenby, and McCulloch (2005) for more details.

[12] Interested readers can find the details of the implementation of the MCMC draws in Chapter 5 of Rossi, Allenby, and McCulloch (2005).

The conditional posterior in (8) is proportional to the product of the likelihood in (7) and the prior of the hyperparameters. I use the Random-Walk Metropolis to obtain the draws of Θ_i. The draws of the hyperparameters in (9) can be broken down into a succession of conditional draws as follows.

$$ind|\gamma, Z, \triangle, \{\mu_k, \Sigma_k\}, \{\Theta\} \tag{10}$$

$$\gamma|ind \tag{11}$$

$$\{\mu_k, \Sigma_k\}|ind, \Theta \tag{12}$$

$$\triangle|ind, Z, \{\mu_k, \Sigma_k\}, \Theta, \tag{13}$$

where the draw of the indicators in (10) is a multinomial draw based on the likelihood ratios with γ_k as the prior probability of membership in each component. The draw of γ given ind in (11) is a Dirichlet draw. The draw of each (μ_k, Σ_k) in (12) can be made using a standard algorithm to draw from a multivariate regression model. The draw of $\triangle$ in (13) requires that the pooling of data from all K components into one regression model.

5.2   *Estimates*

Table 11: Heterogeneity Improves Fit

	Log Marginal Density
Homogenous Coef.	–15786.634
1 Component	–11870.235
5 Components	–10960.389

Table 11 reports the log marginal density for alternative model specifications. The model's posterior probability is monotone in the log marginal density. Hence, a higher log marginal density means a better fit. Note also that log marginal density includes an automatic penalty for the addition of additional parameters (Rossi, Allenby, and McCulloch (2005)). Heterogeneity leads to a substantial improvement in fit. In addition, a more flexible distribution of parameters fits the data better. The estimates from the five-component mixture model yield a higher log marginal density than those from the one-component model.

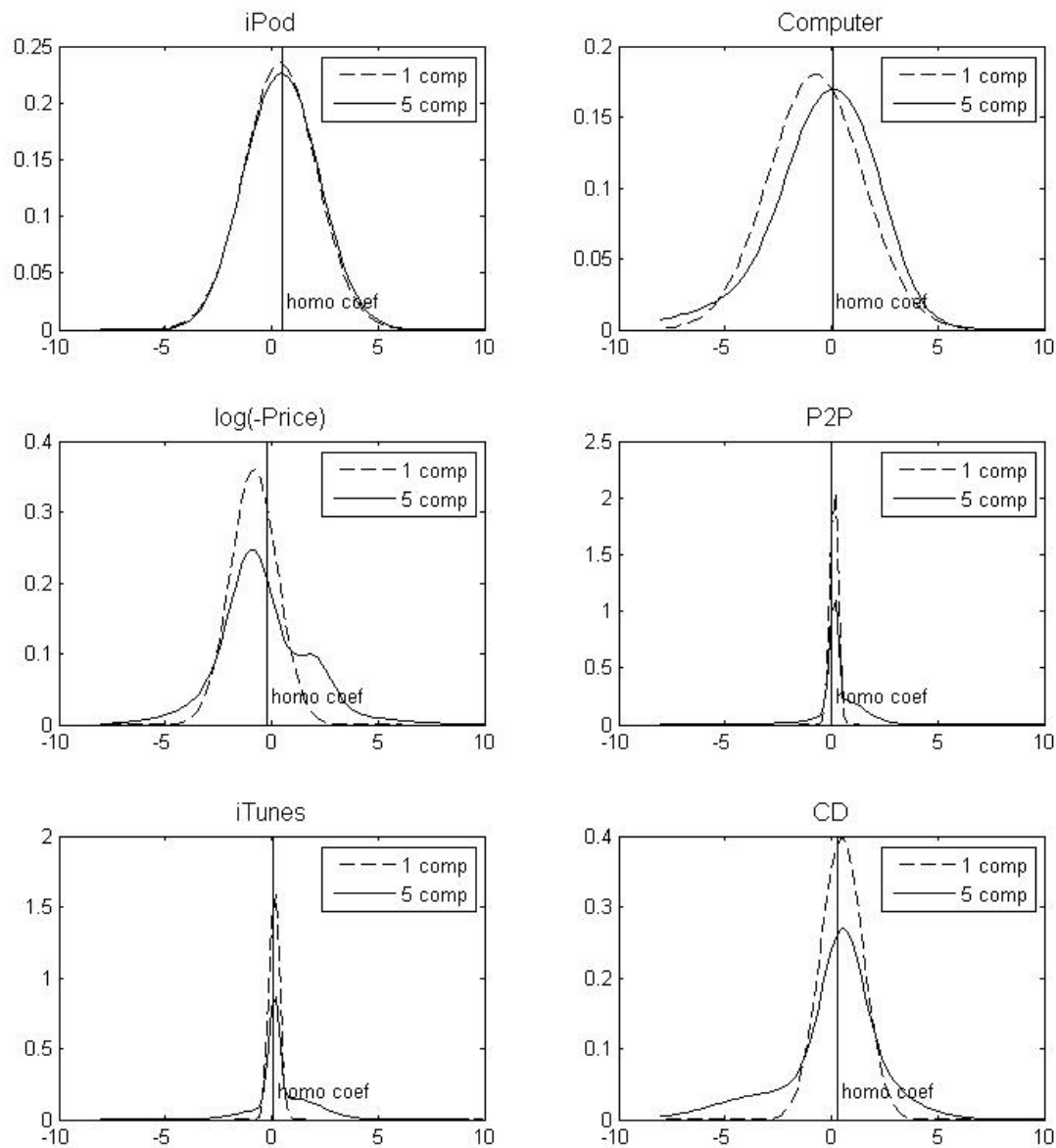Figure 5 displays the density distribution of the coefficients, which indicates substantial

Figure 5:
Density of Random Coefficients

heterogeneity among students' preferences. The five-component model is more flexible and is able to capture the fatter tail of the preference distribution of some of the coefficients.

Table 12: Elasticity of iPod Demand

	1 Component	5 Components
P_{iPod}	-0.229	-0.242
	[-0.251, -0.208]	[-0.283, -0.203]
π	-0.017	-0.017
	[-0.018, -0.016]	[-0.020, -0.015]
f	-0.021	-0.022
	[-0.022, -0.019]	[-0.024, -0.019]
P_{iTunes}	0.007	0.008
	[-0.006, 0.008]	[-0.006, 0.009]
P_{cd}	-0.013	-0.010
	[-0.016, -0.009]	[-0.013, -0.007]

The 5th and 95th percentiles of the estimates are reported in brackets.

Students' demand for iPod is not elastic. Table 12 shows that iPod's own price elasticity ranges between -0.229 and -0.242 under different specifications. Music consumption and the iPod are complements to each other: iPod demand decreases if music prices increase. The elasticities of iPod demand with respect to the prices of music from different sources are all below -0.01%.

6   Counterfactual Experiments

The results in Section 4 suggest that music piracy hurts record sales, and those in Section 5 suggest that it increases demand for, and thus the revenue from, iPods. In this section, I discuss how demand for music and iPods would change in three different copyright regimes, which may shed light on the direction of future copyright policies.

The three copyright regimes are as follows.

Current Regime: In this regime, the RIAA uses the No Electronic Theft Act to occasionally file lawsuits against P2P software companies and individual music pirates. In the first decade of this century, two of the largest P2P software companies, Napster and Kazaa, were sued and later forced to shut down. Between September 2003 and June 2005, 11,700 music pirates were sued (Associated Press (2005)). This wave of lawsuits, however, turned out to have no lasting effect. After a slight decrease in music piracy immediately following the rulings, the number of music pirates began to grow again.

No Music Piracy Regime: In this regime, the government would increase the penalties for music piracy with the aim of eradicating it, which would provide music producers with sufficient income and incentive to create music. Eradicating music piracy would be difficult, but it would be possible if Internet service providers cooperated. Currently there are proposals in France and Britain urging Internet service providers to band together voluntarily to crack down on pirate subscribers.

Free Music Regime: In this regime, music piracy would be legal, and music from all sources would be free. The idea of legalizing music piracy is not new. Fisher (2004) and Netanel (2003) propose a copyright regime in which it is legal to share music files, but firms such as Apple, which would benefit from the boost in iPod sales, would have to pay royalties to the music industry. Such royalties are similar to the private copying levy, a government-mandated scheme in which a levy is charged on transactions involving recordable media. In the U.S., the levy applies to stand-alone CD recorders and portable satellite radio recording devices. Analyzing the demand for music and the demand for iPods can shed light on the range of possible royalties to be imposed on iPods and other MP3 players.

Table 13: Product Attributes in All Regimes

Product Attributes	Current Regime	No Music Piracy Regime	Free Music-Royalty Regime
Price per iPod	$200	$200	$200
Price per iTunes song	$0.99	$0.99	$0
Price per CD	$15	$15	$0
π (in %)	0.01	100	0
Fine per song	$30	$10,000	$0

Table 13 gives the product attributes in the three regimes. The Current Regime describes the current music world. I mimic the current copyright system of the U.S. government and the RIAA's approach with a low probability of getting caught and a small fine for pirating music. At the time the survey was conducted, an 8-gigabyte iPod cost $200, and iTunes songs and CDs cost $0.99 and $15, respectively.

The government and the RIAA would impose a more severe penalty on music piracy in the No Music Piracy Regime. Those engaging in it would be caught for sure, and they would have to pay $10,000 for every song they illegally downloaded.

In the Free Music Regime, downloading music online would free and legal and CDs would

24

cost \$5 to cover the marginal cost of production (I varied this amount from \$1 to \$7, but my main conclusion remained unchanged).

Table 14: Changes in Monthly Music Consumption from Current Regime

To	Free Music Regime	No Music Piracy Regime
P2P	↑ 690%	↓ 100%
iTunes	↑ 70%	↑ 42%
CD	↑ 29%	↑ 24%
Overall Songs	↑ 527%	↓ 68%

In calculating the changes in overall song consumption, I assumed each CD contained 10 songs.

Table 14 reports the changes in the consumption of music from the Current Regime to the two other regimes. When online music becomes free and CDs are sold at a marginal cost in the Free Music Regime, the overall consumption of music increases by more than 500%. When there is no music piracy, the consumption of legally purchased increases, with demand for iTunes songs, and CDs increasing by 42% and 24%, respectively. In other words, music piracy reduces record sales by 24% to 42%. However, because the average student pirates more music than he would consume, the overall consumption of music decreases by 68%.

Table 14 sheds light on how students' probability of purchasing an iPod varies across the regimes. In the Free Music Regime, in which the demand for music is much higher than that in the Current Regime, we should see an increase in the demand for iPods. In the No Music Piracy Regime, in which overall music consumption drops by 68%, we should see a drop in iPod demand.

Table 15: iPod Purchase Probability in Different Regimes (in %)

	Current Regime	Free Music Regime	No Music Piracy Regime
Predicted Market Share	53.04	70.66	46.75
	[51.58, 54.45]	[69.49, 71.78]	[45.15, 48.39]

The 5th and 95th percentiles of the estimates are reported in brackets.

Table 15 shows students' iPod purchase probabilities are consistent with the changes in music consumption in the different regimes. On average, students would be approximately 18% more likely to buy an iPod in the Free Music Regime than in the Current Regime because of the increase in overall music consumption. That would translate into a \$36 increase in expected revenue from each student.

When the average student's overall music consumption decreases in the No Music Piracy Regime, he becomes approximately 6% less likely to buy an iPod. Put another way, in this scenario, approximately 12% of Apple's revenue from iPods comes from the additional music consumption realized by music piracy. Using Apple's iPod revenue in 2008 ($9.2 billion) as a reference, we can estimate that $1.1 billion of the company's revenue that year came from music piracy.

7 Conclusion

I answer two important empirical questions regarding music piracy in this paper. First, does music piracy hurt record sales, and, if so, by how much? Second, does music piracy benefit the sales of music complements, such as iPods and, if so, by how much? I answer these questions using a unique conjoint survey data set. Estimates from three-stage least-squares estimation indicate that music piracy does indeed hurt record sales, which corroborates the prevailing belief in the music industry, but is contrary to the claim of Oberholzer-Gee and Strumpf (2007). I also use a Bayesian approach to estimate the demand for iPods and find that music piracy benefits iPod sales. If there were no music piracy, Apple's revenue from iPod sales would decrease by 12%, which would have translated to $1.1 billion in 2008. If music were free, however, then Apple's expected revenue would increase by $36 per student.

In addition to iPods, many other products are also music complements and thus would benefit from a free music regime. Examples include other brands of MP3 players, Internet providers, and live music performances. Although I only focus on iPods in this paper, my approach could easily be extended to examine the complementary relationships between music and these other products. This extension would make possible the evaluation of the effects of different copyright regimes on different products.

References

ASSOCIATED PRESS (2005): "Rethinking the File-Swap Morass," `http://www.wired.com/entertainment/music/news/2005/06/68000`.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

BLACKBURN, D. (2004): "Does File Sharing Affect Record Sales?," Ph.D. thesis, Harvard University.

CARLSSON, F., AND P. MARTINSSON (2001): "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments?," *Journal of Environmental Economics and Management*, 41(2), 179–192.

FISHER, W. W. (2004): *Promises to Keep*. Stanford University Press, Stanford, CA.

GREEN, P. E., A. M. KRIEGER, AND Y. J. WIND (2004): "Adventures in Conjoint Analysis: A Practitioner's Guide to Trade-Off Modeling and Applications," `http://marketing.wharton.upenn.edu/people/faculty/green/green_monograph.cfm`.

GREEN, P. E., AND V. R. RAO (1971): "Conjoint Measurement for Quantifying Judgment Data," *Journal of Marketing Research*, 8(3), 355–363.

HENSHER, D. (1994): "Stated Preference Analysis of Travel Choices: The State of Practice," *Transportation*, 21(2), 107–133.

HENSHER, D., AND J. LOUVIERE (1983): "Using Discrete Choice Models with Experimental Design Data to Forecast Consumer Demand for a Unique Cultural Event," *The Journal of Cultural Research*, 10(3), 348–361.

HENSHER, D., J. LOUVIERE, AND J. SWAIT (1999): "Combining Sources of Preference Data," *Journal of Econometrics*, 89(1-2), 197–221.

LEUNG, T. C. (2012): "What is the True Loss Due to Piracy? Evidence from Microsoft Office in HK," *Review of Economics and Statistics*, forthcoming.

LIEBOWITZ, S. J. (2006): "File Sharing: Creative Destruction or Just Plain Destruction?," *Journal of Law and Economics*, 49(1), 1–28.

——— (2007): "How Reliable is the Oberholzer-Gee and Strumpf Paper on File Sharing?," http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1014399.

LOUVIERE, J., AND G. WOODWORTH (1983): "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20(4), 350–367.

MORTIMER, J. H., C. NOSKO, AND A. SORENSEN (2012): "Supply responese to digital distribution: Recorded music and live performances," *Information Economics and Policy*, 24(1), 3–14.

NETANEL, N. (2003): "Impose a Noncommercial Use Levy to Allow Free Peer-to-Peer File Sharing," *The University of Texas School of Law*, (009), Law and Economics Working Paper.

NEVO, A. (2000): "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4), 513–548.

OBERHOLZER-GEE, F., AND K. STRUMPF (2007): "Effect of File Sharing on Record Sales: An Empirical Analysis," *Journal of Political Economy*, 115(1), 1–42.

——— (2010): "File Sharing and Copyright," in *NBER Book Series Innovation Policy and the Economy*, vol. 10.

ORME, B. K. (2005): *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Research Publisher LLC, Madison, WI.

PEITZ, M., AND P. WAELBROECK (2004): "The Effect of Internet Piracy on Music Sales: Cross-Section Evidence," *Review of Economics Research on Copyright Issues*, 1(2), 71–79.

PETRIN, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4), 705–729.

Rob, R., and J. Waldfogel (2006): "Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students," *Journal of Law and Economics*, 49(1), 29–62.

Rossi, P., G. Allenby, and R. McCulloch (2005): *Bayesian Statistics and Marketing*. John Wiley and Sons, Hoboken, NJ.

Sawtooth Software (2008): "The CBC System for Choice-Based Conjoint Analysis," *Sawtooth Software Technical Paper Series*.

Shiller, B., and J. Waldfogel (2011): "Music for a Song: An Emirical Look at Uniform Song Pricing and its Alternatives," *Journal of Industrial Economics*, 59(4), 630–660.

Zentner, A. (2006): "Measuring the Effect of File Sharing on Music Purchases," *Journal of Law and Economics*, 49(1), 63–90.